

PRE-ANALYSIS PLAN FOR ROUND 4 SCHOOLING, INCOME, AND HEALTH RISK IN MALAWI (SIHR) DATA

Principal Investigators:

Sarah Baird, University of Otago
Ephraim Chirwa, Chancellor College
Craig McIntosh, UCSD
Berk Özler, World Bank and University of Otago

Analysis plan:

The core analysis will compare the impact of the Conditional Cash Transfer (CCT) and Unconditional Cash Transfer (UCT) treatment to control EAs for the baseline schoolgirl stratum, and will compare the CCT treatment to the control for the baseline dropout stratum. Most of the analysis will consist of Round 4 cross-sectional regression (using OLS unless not appropriate), although where possible we will also pursue panel difference-in-differences analysis for variables that have been consistently collected in multiple rounds. For consistency, the analysis will include the full set of controls used in the paper Baird, McIntosh and Özler (2011). These controls include baseline values of the following: a household asset index, highest grade attended, a dummy variable for having started sexual activity, dummy variables for age, and strata dummies. Standard errors will be clustered at the EA level, and results will be weighted to make them representative of the target population in the study EAs.

Only if a significant impact is found in the core analysis will further heterogeneity of impact be explored. Heterogeneity will be explored along both experimental dimensions including the amount of the transfer and the split between the parent and the girl, as well as based on the age at which you got the program and differences between rural and urban.

Note on construction of indexes:

To construct indexes for classes of variables, we will adhere to the following rubric:

- a) For each sub-question in a family of variables, first align answers so that higher numbers always have a consistent meaning (good or bad).
- b) Calculate the mean and SD of the responses to each sub-question in the sample *in the control group – separately for baseline schoolgirls and baseline dropouts*.
- c) Create normalized variables that have the mean subtracted off and are divided by the SD.
- d) Calculate the raw mean of the normalized variables for all sub-questions within a family of variables. This mean is the ‘index’ for those variables. *This summary index can further be normalized if desired.*

For the core analysis we will not pursue the analysis of sub-variables within an index unless the index as a whole is significant.

Construction of Core Indexes: Primary outcomes are indicated in bold text.

1. Core Respondent-level outcomes: Can be analysis with simple cross-sections comparing UCT, CCT, and control. No extensive margin issue with any of these variables.

- a. Schooling and Marriage (replication of QJE results with age-appropriate dependent variables):
 - i. **Highest grade completed (S7, Q7)**
 - ii. Highest educational qualification achieved (S7, Q9)
 - iii. **Achievement, replacing the test scores with the ‘competencies’ (see Appendix A1 for construction). We will show the components, as well as the index for the quality index; and show only an index for the quantity.**
 - iv. **Ever married (Part II CS, Q2e), ever pregnant (S18, Q1, Q2), number of live births (S18, Q17).**
 - v. **Hazard model of age of first marriage (S14, Q1 (and Round 3 data for those already married at Round 3)) and age at first birth (construct using age of respondent and DOB), with ‘uncompleted spells’ for those never married or never first birth.**
 - vi. Sexual behavior: ABC, # partners ever, as in previous papers. Ever had sex (S12, Q2, Q3, Q4) age at first sex (S12 Q4), total number of partners ever (S12, Q5), sexually active in past 12 months (S12, Q7), condom use last sex with most recent partner (S12, Q23), having a partner with an age difference of more than 5 years (S12Q12)

- b. Health: Replication of Lancet results (Baird, Garfein, McIntosh and Özler)
 - i. **HIV prevalence in R4, HIV incidence R3-R4.**
 - ii. **Anemia: construct binary measure based on definition of mild anemia from WHO and used in survey (with different thresholds if pregnant), (with bednets, breastfeeding, timing of last meal, taking medical for anemia, and menstruating as secondary moderator channel) (VCT, S1).**
 - iii. Use of reliable birth control: S12 Q27 anything except withdrawal, periodic abstinence, other
 - iv. Desired fertility S16 Q4 or Q10.
 - v. Mental health, calculated as in Baird, de Hoop, Ozler (2013) (S9, Q9-20): binary
 - vi. Number of meals eaten with meat, eggs, fish in past 7 days (S9, Q 6-8)

- c. Empowerment & aspirations:
 - i. Index of self-efficacy: S11a Q1-10.
 - ii. Index of preferences for child education: S11a Q17-25.
 - iii. Index of social participation: S11a Q13,14,16.
 - iv. Aspirations: Change in ladder from five years ago to five years from now (S9, Q23-Q21)
 - v. **Change in ladder from five years ago to today (S9, Q22-Q21)**

Super-index of overall empowerment i-iv.

- d. Wages/Employment:
 - i. Effective wage: S8 Q2 d-g, Q6/Q4 – converted into a daily wage rate.
 - ii. **Opportunity cost of time: Minimum wage from S8 Q13-17.**
 - iii. Labor income: S8 Q7.
 - iv. **Typical wage: S8 Q10.**
 - v. Any wage work in past 3 months (S8 Q9)
 - vi. **Sector of employment: S8: Sum of Q4 a&c divided by the whole sum of Q4.**

- e. Consumption:

- i. Impact on household-level consumption aggregate (Real Total comp. monthly consumption in market unit prices in USD per person), constructed as in previous rounds
2. Married Core Respondent Outcomes: to be analyzed within the sample of married core respondents, contextualized by impacts on marriage rates and age at first marriage but no attempt at a 'Heckman-style' correction. Secondary analysis will delve more into intensive/extensive margin impacts as necessary.
- a. Core Respondent's Empowerment married:
 - i. Index of financial decisionmaking: S15 Q1-5, 'Resp'=2. 'Joint'=1, other=0¹
 - ii. Index of marital satisfaction: S15 Q20-29.
 - iii. Index of women's divorce prospects: S15 Q39-40, S15 Q48a-d
 - iv. Index of fertility disempowerment: S16 Q6==3 & Q7==(1|2), Q39-40.
 - v. Index of self-determination in marriage: S16 Q20-28
 - vi. Index of frequency of social contact: S16 Q29+30+32
 - vii. Index of spousal abuse: S16 Q41-46, 47, 50, 52
 - viii. Age difference between wife and husband (S12, Q12-Part II CS, Q2c)
 - ix. Female ag decisionmaking power: If the answer to any of the questions S3 Q12, Q16, or Q17 is the "CR" for any plot, the dummy variable for the entire HH is defined as a '1' and '0' otherwise.
 - x. Female microenterprise participation: S6 Q11 Core Resp controlling any enterprise profits
 - xi. Female livestock control: S3 Q22: Core Resp responsible for any livestock decisionmaking.
 - xii. Ratio of female to male-specific consumption: calculate total spending for CR (per month or per 12 months) on all items asked and divide it by the same variable calculated for the husband from Section 27 (S10 Q3,4,7,8 normalized sum) / (S27 Q3,4,7,8 normalized sum).

Super-index of empowerment: i-viii.

Super-index of economic control: iv-xii.

- b. Husband Quality.
 - i. Husband's highest grade completed, highest certificate attained. S25 Q2,4
 - ii. Husband's wage rate S26 Q5
 - iii. Currently employed S26 Q6..
 - iv. Husband's score on cognitive test
 - v. Husband HIV status.
 - vi. Husband marital fidelity. Partners ever: S32 Q2, Partners 12 mo. S32 Q3. Concurrence: S32 Q15 answer for spouse (column 1)
 - vii. Husband's mental health (constructed in same manner as CR) and then standardized.

Super-index of husband quality: i-vii.

- c. Husband Gender Empowerment.
 - i. Husband Index of GE: S30 Q1-9
 - ii. Husband Index of wife's autonomy: S30 Q13-21.

¹ Also, as a rule, indices for each individual include the items with no missing values. We do not impute values for item non-response, or exclude variables or individuals

- iii. Husband Index of justification for abuse S30 Q26-28.
- iv. Husband Index of divorce prospects: S31 Q10-11, 15,16
- v. Husband desired fertility. S30 Q49

Super-index of husband gender empowerment: i-v.

3. Unmarried Core Respondent Outcomes: to be analyzed within the sample of unmarried core respondents.

- a. Empowerment unmarried:
 - i. Index of autonomy: S17 Q2a-d, Q4, Q6
 - ii. Index of abuse: S17 Q14,16,17a-e.

Super-index of unmarried empowerment: i-ii.

4. Child-Level Outcomes: Can be analyzed two different ways. First, unconditionally examining simple comparison between treatment and control. Secondly, conditionally including dummies for the age of child and the age of the mother so as to drop out effects coming from delayed fertility.

- a. Health Outcomes:
 - i. Birthweight: S21 Q20
 - ii. Vaccinations: S22 Q5, 7, 8, 10, 11 (index)
 - iii. Neonatal/Infant/Child mortality: S18 Q16-21; S21 Q10 (use Round 3 data for children who had already passed away by Round 3)**
 - iv. Bednets: S22 Q 13
- b. Parental Practices:
 - i. **Breastfeeding (6 months exclusive or to death is died younger than 6 months): S21 Q22, Q23, Q25**
 - ii. **Parenting: S23 Q3-16, 21, 24, 27 (index)**
- c. Anthropometrics:
 - i. **Height for age z-score, weight for height z-score, nutritional status (based on weight for height using Malawi standards)**
 - ii. Stunting, wasting (binary, less than 2 std z-score)
- d. Educational Testing:
 - i. **MDAT scores**
 - ii. **SDQ scores on behavior towards others.** Use the following to score:
<http://www.sdqinfo.com/c3.html>

Appendix 1: Construction of Competencies Index:

Moderator variable for fertilizer application: S11bQ22

Fertilizer (Q23-26):

Quantity index: time taken to complete (Q23), categorize it as 1 'below median' (in seconds); 2 'above median'; and 3 'did not complete/did not complete in time'. Median is calculated among those who completed under the allocated time.

Quality index: Each Q (24-26) coded as 1 if Yes 0 if No and then added up to create an index between 0-3 of the quality of the application of fertilizer.

Normalize each by subtracting the control mean and dividing by the control SD.

Making change (Q27-28):

Same as above: (quantity index, Q27) and quality index (Q28).

Use the same procedure for Q29-30, Q31-32. Then, add the quality indices (Q28, 30, and 32). Add quantity indices (Q27, 29, and 31).

Normalize each by subtracting the control mean and dividing by the control SD.

Sending a text message (Q35-37) – moderator variables to be used for adjustment (Q33-34):

Same as above, then normalize each index.

Use the calculator on mobile phone (Q38-39):

Same as above, then normalize each index.

Calculate profits from trade (Q-40-42):

Same as above, then normalize each index.

Finally, average the normalized quantity indices and the quality indices separately to produce two final competency indices.